Translating basic science discoveries to clinical practice—Let us not repeat the naiveté of the pre-omics era

Kimberly J. Payne[1], Abigail Benitez[2], Sinisa Dovat[3]

[1]Department of Pathology and Human Anatomy, Loma Linda University, Loma Linda, CA 92350, USA; [2]Transplant Institute, Loma Linda University Medical Center, CA 92354, USA; [3]Department of Pediatrics, Pennsylvania State University Medical College, Hershey, PA 17033, USA
Correspondence to: Kimberly J. Payne. Loma Linda University, 11085 Campus Street, Loma Linda, CA 92350, USA. Email: kpayne@llu.edu.

The genomics era has produced large data sets that have proved more difficult than anticipated to integrate into basic science/translational research, and even more so in clinical practice. Part of the challenge is that we saw the promise of more information without perceiving that this new data would not simply be applied or used in the same manner as in the past. Big Data has pushed us to develop strategies for integrating the classic bench research with information from bioinformatics and whole genome data in ways that will allow us to identify targets, develop and test therapeutic strategies, and integrate these strategies into general clinical practice. Successful translational strategies are emerging (1) with some aspects that are broadly applicable, although each clinical problem has its own challenges. Currently emerging strategies provide a foundation for what must continue to be a rapidly evolving process.

One such translational strategy, described by Bruhn *et al*. (1), facilitated identification and validation of a diagnostic and therapeutic target gene in allergy. A module-based approach was used to prioritize candidate genes from among the large numbers of genes up or down regulated in this disease. The strategy was based on the premise that the most relevant disease-associated genes are components of networks or modules of genes that are functionally related. Genomic and bioinformatic data were integrated with functional approaches that incorporated animal models and patient samples to provide translational data relevant to both diagnosis and therapy in allergy. Aspects of module-based strategies such as this may prove generalizable to a broad range of translational studies, provided that the approach is tailored to meet unique challenges of specific clinical problems and to capitalize on novel disease-specific assays

and animal models, as well as available clinical data and patient samples. The process of identifying such genes can be hampered by disease heterogeneity as well as differences in protein-protein interactions (PPI) among gene networks or module genes in different cell types. Bruhn *et al*. used seasonal allergic rhinitis (SAR) as a disease model because it is has a well-defined phenotype and pathogenesis–not all diseases are so obliging.

Allergy as a disease model offered several unique resources that could be exploited to maximize the information gained. Peripheral blood mononuclear cells provided an easily obtainable source of human CD4+ T cells, key players in disease pathogenesis. CD4+ T cells (I) provided an *in vitro* assay of response to pollen challenge; (II) could be easily evaluated by microarray and (III) provided a human counterpart for *in vivo* functional studies in mouse models of allergy. Mouse models of allergy are well established and could be tested in gene knockout mice. Nasal fluid from patients in hay fever season offered an easily obtainable human sample for evaluating gene candidates as diagnostic markers—provided the gene product was a secreted protein.

To define a gene module relevant for SAR and other allergies, Bruhn *et al*. sought to identify a co-regulated group of genes that included IL-13, a key cytokine in multiple components of the allergic response. Module selection was based on gene network studies which showed that genes acting together in a specific disease are often regulated by the same transcription factors (2-5). Known transcription factors that regulate IL-13 were identified from literature reports and additional candidates were identified based on bioinformatics predictions from

microarray analysis of patient samples stimulated *in vitro*. Seven transcription factors were verified as regulators of IL-13 by siRNA knockdown which also identified co-regulated genes. Genes in a module co-regulated with IL-13 were also shown to be close interactors using the human PPI network. From these, a gene *(S100A4)* producing a highly expressed secreted protein was selected as a candidate diagnostic marker likely to be present in body fluids.

The *S100A4* gene was validated as a diagnostic marker and a potential therapeutic target in allergy using multiple models. S100A4 was shown to be upregulated in nasal fluids of allergy patients and skin biopsies from patients with allergic dermatitis. *In vitro* and *in vivo* functional assays validated S100A4 as a target for therapy. These included S100A4 protein-blocking assays performed on cultures of antigen-challenged human CD4 T cells and *S100A4* deletion in a mouse model of allergy. *In vivo* studies were facilitated by the availability of *S100A4* gene knockout mice (6) that had previously been studied in context of cancer (7).

This study by Bruhn *et al.* along with the work of other groups on leukemia and neuroblastoma (8-13) hint at the fulfillment of the promise we anticipated with the advent of the omics era. However, a number of challenges remain. The road from the moment the target is identified—in this case S100A4—until the correct drug or antibody is developed, patented, tested in multiple clinical trials and finally approved for wide clinical use is very long and very expensive. Streamlining regulatory processes and producing a flexible infrastructure that is responsive to the increasingly rapid rates of change in the drug discovery process while maintaining patient protection is critical to fulfilling the promise of current translational strategies.

Ensuring that the long term medical benefits are realized from current translational studies will also depend on developing an educational infrastructure that can meet the needs of the emerging research community. Training clinical and translational researchers in good mentoring strategies (14) develops skills that remain foundational as the processes of translational research changes. Team approaches that enhance the interface between bioinformatics, basic bench research and translational/clinical studies will be essential. Group learning activities (preferably real world) that challenge graduate and medical students to develop translational research strategies to address clinical challenges in a research team context will provide an opportunity to develop unique skills. These specialized skills will be essential for maintaining and enhancing the translational research workforce needed to develop the therapies that can come from translational research strategies such as that described by Bruhn *et al*.

In our naiveté many of us imagined that omics would simply allow us to do what we were already doing better and faster. However Big Data created a "push" that fundamentally changed how we approach biomedical research. This emergent process is continuing to unfold and reaping both the short- and long-term rewards will require the development of regulatory and educational infrastructures nimble enough to adapt to the emergent landscape of translational research.

Acknowledgements

References

1. Bruhn S, Fang Y, Barrenäs F, et al. A generally applicable translational strategy identifies S100A4 as a candidate gene in allergy. Sci Transl Med 2014;6:218ra4.
2. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. Cell 2011;144:986-98.
3. Ideker T, Krogan NJ. Differential network biology. Mol Syst Biol 2012;8:565.
4. Amit I, Regev A, Hacohen N. Strategies to discover regulatory circuits of the mammalian immune system. Nat Rev Immunol 2011;11:873-80.
5. Ideker T, Sharan R. Protein networks in disease. Genome Res 2008;18:644-52.
6. Li ZH, Dulyaninova NG, House RP, et al. S100A4 regulates macrophage chemotaxis. Mol Biol Cell 2010;21:2598-610.
7. Malashkevich VN, Dulyaninova NG, Ramagopal UA, et al. Phenothiazines inhibit S100A4 function by inducing protein oligomerization. Proc Natl Acad Sci U S A

2010;107:8605-10.

8. Daigle SR, Olhava EJ, Therkelsen CA, et al. Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. Cancer Cell 2011;20:53-65.

9. Bernt KM, Zhu N, Sinha AU, et al. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. Cancer Cell 2011;20:66-78.

10. Krivtsov AV, Feng Z, Lemieux ME, et al. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14:355-68.

11. Zuber J, Shi J, Wang E, et al. RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature 2011;478:524-8.

12. Bresler SC, Wood AC, Haglund EA, et al. Differential inhibitor sensitivity of anaplastic lymphoma kinase variants found in neuroblastoma. Sci Transl Med 2011;3:108ra114.

13. George RE, Attiyeh EF, Li S, et al. Genome-wide analysis of neuroblastomas using high-density single nucleotide polymorphism arrays. PLoS One 2007;2:e255.

14. Pfund C, House S, Spencer K, et al. A research mentor training curriculum for clinical and translational researchers. Clin Transl Sci 2013;6:26-33.